

EFFICIENT METHODS AND APPARATUS FOR HIGH-THROUGHPUT PROCESSING OF GENE SEQUENCE DATA

5

ABSTRACT OF THE DISCLOSURE

One disclosed method of processing gene sequence data includes the steps of reading gene sequence data corresponding to a gene sequence and coding sequence data corresponding to a plurality of coding sequences within the gene sequence; identifying and storing, by following a set of primer selection rules, primer pair data within the gene sequence data for one of the coding sequences; repeating the acts of identifying and storing such that primer pair data are obtained for each sequence of the plurality of coding sequences; and simultaneously amplifying the plurality of coding sequences in gene sequences from three or more of individuals using the identified pairs of primer sequences. The set of primer selection rules include a rule specifying that all of the primer pair data for the plurality of coding sequences be obtained for a predetermined annealing temperature, which allows for the subsequent simultaneous amplification of sequences from hundreds of individuals in a single amplification run.